

S. GUCKER

RE-RUN



OIPE

#34

RAW SEQUENCE LISTING

PATENT APPLICATION: US/08/879,469

DATE: 11/25/2003

TIME: 07:40:36

Input Set : N:\Crf3\RULE60\08879469.RAW.txt

Output Set: N:\CRF4\11252003\H879469.raw

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Murgita, Robert A.

7 (ii) TITLE OF INVENTION: RECOMBINANT HUMAN ALPHA-FETOPROTEIN AS A

8 CELL PROLIFERATIVE AGENT

10 (iii) NUMBER OF SEQUENCES: 16

12 (iv) CORRESPONDENCE ADDRESS:

13 (A) ADDRESSEE: Fish & Richardson P.C.

14 (B) STREET: 225 Franklin Street, Suite 3100

15 (C) CITY: Boston

16 (D) STATE: MA

17 (E) COUNTRY: USA

18 (F) ZIP: 02110-2804

20 (v) COMPUTER READABLE FORM:

21 (A) MEDIUM TYPE: Floppy disk

22 (B) COMPUTER: IBM PC compatible

23 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

24 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

26 (vi) CURRENT APPLICATION DATA:

C--> 27 (A) APPLICATION NUMBER: US/08/879,469

C--> 28 (B) FILING DATE: 20-Jun-1997

W--> 33 (C) CLASSIFICATION: 435

30 (vii) PRIOR APPLICATION DATA:

31 (A) APPLICATION NUMBER: US/08/377,316

32 (B) FILING DATE: 24-JAN-1995

36 (viii) ATTORNEY/AGENT INFORMATION:

37 (A) NAME: Clark, Paul T.

38 (B) REGISTRATION NUMBER: 30,162

39 (C) REFERENCE/DOCKET NUMBER: 06727/006001

41 (ix) TELECOMMUNICATION INFORMATION:

42 (A) TELEPHONE: (617) 542-5070

43 (B) TELEFAX: (617) 542-8906

44 (C) TELEX: 200154

47 (2) INFORMATION FOR SEQ ID NO: 1:

49 (i) SEQUENCE CHARACTERISTICS:

50 (A) LENGTH: 2022 base pairs

51 (B) TYPE: nucleic acid

52 (C) STRANDEDNESS: single

53 (D) TOPOLOGY: linear

W--> 55 (ii) MOLECULE TYPE: DNA

60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

62 ATATTGTGCT TCCACCACTG CCAATAACAA AATAACTAGC AACCATGAAG TGGGTGGAAT 60

64 CAATTTTTTT AATTTTCCTA CTAAATTTTA CTGAATCCAG AACACTGCAT AGAAATGAAT 120

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/08/879,469

DATE: 11/25/2003

TIME: 07:40:36

Input Set : N:\Crif3\RULE60\08879469.RAW.txt

Output Set: N:\CRF4\11252003\H879469.raw

```

66 ATGGAATAGC TTCCATATTG GATTCTTACC AATGTACTGC AGAGATAAGT TTAGCTGACC 180
68 TGGCTACCAT ATTTTTTGCC CAGTTTGTTT AAGAAGCCAC TTACAAGGAA GTAAGCAAAA 240
70 TGGTGAAAGA TGCATTGACT GCAATTGAGA AACCCACTGG AGATGAACAG TCTTCAGGGT 300
72 GTTTAGAAAA CCAGCTACCT GCCTTTCTGG AAGAACTTTG CCATGAGAAA GAAATTTTGG 360
74 AGAAGTACGG ACATTGAGAC TGCTGCAGCC AAAGTGAAGA GGAAGACAT AACTGTTTTT 420
76 TTGCACACAA AAAGCCCCTT GCAGCATGGA TCCCACCTTTT CCAAGTTCCA GAACCTGTCA 480
78 CAAGCTGTGA AGCATATGAA GAAGACAGGG AGACATTCAT GAACAAATTC ATTTATGAGA 540
80 TAGCAAGAAG GCATCCCTTC CTGTATGCAC CTACAATTCT TCTTTCGGCT GCTGGGTATG 600
82 AGAAAATAAT TCCATCTTGC TGCAAAGCTG AAAATGCAGT TGAATGCTTC CAAACAAAGG 660
84 CAGCAACAGT TACAAAAGAA TTAAGAGAAA GCAGCTTGTT AAATCAACAT GCATGTCCAG 720
86 TAATGAAAAA TTTTGGGACC CGAACTTTCC AAGCCATAAC TGTTACTAAA CTGAGTCAGA 780
88 AGTTTACCAA AGTTAATTTT ACTGAAATCC AGAACTAGT CCTGGATGTG GCCCATGTAC 840
90 ATGAGCACTG TTGCAGAGCA GATGTGCTGG ATTGTCTGCA GGATGGGGAA AAAATCATGT 900
92 CCTACATATG TTCTCAACAA GACACTCTGT CAAACAAAAT AACAGAATGC TGCAAAGTGA 960
94 CCACGCTGGA ACGTGGTCAA TGTATAATTC ATGCAGAAAA TGATGAAAAA CCTGAAGGTC 1020
96 TATCTCCAAA TCTAAACAGG TTTTTAGGAG ATAGAGATTT TAACCAATTT TCTTCAGGGG 1080
98 AAAAAAATAT CTTCTTGGCA AGTTTTGTTC ATGAATATTC AAGAAGACAT CCTCAGCTTG 1140
100 CTGTCTCAGT AATTCTAAGA GTTGCTAAAG GATACCAGGA GTTATTGGAG AAGTGTTTTCC 1200
102 AGACTGAAAA CCCTCTTGAA TGCCAAGATA AAGGAGAAGA AGAATTACAG AAATACATCC 1260
104 AGGAGAGCCA AGCATTGGCA AAGCGAAGCT GCGGCCTCTT CCAGAACTA GGAGAATATT 1320
106 ACTTACAAAA TGAGTTTCTC GTTGCTTACA CAAAGAAAGC CCCCAGCTG ACCTCGTCGG 1380
108 AGCTGATGGC CATCACCAGA AAAATGGCAG CCACAGCAGC CACTTGTTGC CAACTCAGTG 1440
110 AGGACAAACT ATTGGCCTGT GCGGAGGGAG CGGCTGACAT TATTATCGGA CACTTATGTA 1500
112 TCAGACATGA AATGACTCCA GTAAACCCTG GTGTTGGCCA GTGCTGCACT TCTTCATATG 1560
114 CCAACAGGAG GCCATGCTTC AGCAGCTTGG TGGTGGATGA AACATATGTC CCTCCTGCAT 1620
116 TCTCTGATGA CAAGTTCATT TTCCATAAGG ATCTGTGCCA AGCTCAGGGT GTAGCGCTGC 1680
118 AAAGGATGAA GCAAGAGTTT CTCATTAACC TTGTGAAGCA AAAGCCACAA ATAACAGAGG 1740
120 AACAACCTGA GGCTCTCATT GCAGATTTCT CAGGCCTGTT GGAGAAATGC TGCCAAGGCC 1800
122 AGGAACAGGA AGTCTGCTTT GCTGAAGAGG GACAAAAACT GATTTCAAAA ACTGGTGCTG 1860
124 CTTTGGGAGT TTAAATTAAT TCAGGGGAAG AGAAGACAAA ACGAGTCTTT CATTCGGTGT 1920
126 GAACTTTTCT CTTTAATTTT AACTGATTTA ACACTTTTTG TGAATTAATG ATAAAGACTT 1980
128 TTATGTGAGA TTTCTTATC ACAGAAATAA AATATCTCCA AA 2022

```

130 (2) INFORMATION FOR SEQ ID NO: 2:

132 (i) SEQUENCE CHARACTERISTICS:

133 (A) LENGTH: 590 amino acids

134 (B) TYPE: amino acid

135 (C) STRANDEDNESS: Not Relevant

136 (D) TOPOLOGY: linear

138 (ii) MOLECULE TYPE: protein

143 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```

145 Thr Leu His Arg Asn Glu Tyr Gly Ile Ala Ser Ile Leu Asp Ser Tyr
146 1 5 10 15
148 Gln Cys Thr Ala Glu Ile Ser Leu Ala Asp Leu Ala Thr Ile Phe Phe
149 20 25 30
151 Ala Gln Phe Val Gln Glu Ala Thr Tyr Lys Glu Val Ser Lys Met Val
152 35 40 45
154 Lys Asp Ala Leu Thr Ala Ile Glu Lys Pro Thr Gly Asp Glu Gln Ser
155 50 55 60
157 Ser Gly Cys Leu Glu Asn Gln Leu Pro Ala Phe Leu Glu Glu Leu Cys

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/08/879,469

DATE: 11/25/2003

TIME: 07:40:36

Input Set : N:\Crf3\RULE60\08879469.RAW.txt

Output Set: N:\CRF4\11252003\H879469.raw

158	65		70		75		80
160	His	Glu	Lys	Glu	Ile	Leu	Glu
161				85			
163	Gln	Ser	Glu	Glu	Arg	His	Asn
164			100				105
166	Thr	Ala	Ala	Trp	Ile	Pro	Leu
167			115				120
169	Cys	Glu	Ala	Tyr	Glu	Glu	Asp
170		130				135	
172	Tyr	Glu	Ile	Ala	Arg	Arg	His
173		145				150	
175	Leu	Ser	Ala	Ala	Gly	Tyr	Glu
176				165			170
178	Glu	Asn	Ala	Val	Glu	Cys	Phe
179			180				185
181	Glu	Leu	Arg	Glu	Ser	Ser	Leu
182			195				200
184	Lys	Asn	Phe	Gly	Thr	Arg	Thr
185		210				215	
187	Ser	Gln	Lys	Phe	Thr	Lys	Val
188		225				230	
190	Leu	Asp	Val	Ala	His	Val	His
191			245				250
193	Asp	Cys	Leu	Gln	Asp	Gly	Glu
194			260				265
196	Gln	Asp	Thr	Leu	Ser	Asn	Lys
197		275				280	
199	Leu	Glu	Arg	Gly	Gln	Cys	Ile
200		290				295	
202	Glu	Gly	Leu	Ser	Pro	Asn	Leu
203		305				310	
205	Asn	Gln	Phe	Ser	Ser	Gly	Glu
206			325				330
208	His	Glu	Tyr	Ser	Arg	Arg	His
209			340				345
211	Arg	Val	Ala	Lys	Gly	Tyr	Gln
212		355				360	
214	Glu	Asn	Pro	Leu	Glu	Cys	Gln
215		370				375	
217	Tyr	Ile	Gln	Glu	Ser	Gln	Ala
218		385				390	
220	Gln	Lys	Leu	Gly	Glu	Tyr	Tyr
221			405				410
223	Thr	Lys	Lys	Ala	Pro	Gln	Leu
224			420				425
226	Arg	Lys	Met	Ala	Ala	Thr	Ala
227		435				440	
229	Lys	Leu	Leu	Ala	Cys	Gly	Glu
230		450				455	

RAW SEQUENCE LISTING

PATENT APPLICATION: US/08/879,469

DATE: 11/25/2003

TIME: 07:40:36

Input Set : N:\Crif3\RULE60\08879469.RAW.txt

Output Set: N:\CRF4\11252003\H879469.raw

```

232   Leu Cys Ile Arg His Glu Met Thr Pro Val Asn Pro Gly Val Gly Gln
233   465                               470                               475                               480
235   Cys Cys Thr Ser Ser Tyr Ala Asn Arg Arg Pro Cys Phe Ser Ser Leu
236                               485                               490                               495
238   Val Val Asp Glu Thr Tyr Val Pro Pro Ala Phe Ser Asp Asp Lys Phe
239                               500                               505                               510
241   Ile Phe His Lys Asp Leu Cys Gln Ala Gln Gly Val Ala Leu Gln Arg
242                               515                               520                               525
244   Met Lys Gln Glu Phe Leu Ile Asn Leu Val Lys Gln Lys Pro Gln Ile
245                               530                               535                               540
247   Thr Glu Glu Gln Leu Glu Ala Leu Ile Ala Asp Phe Ser Gly Leu Leu
248   545                               550                               555                               560
250   Glu Lys Cys Cys Gln Gly Gln Glu Gln Glu Val Cys Phe Ala Glu Glu
251                               565                               570                               575
253   Gly Gln Lys Leu Ile Ser Lys Thr Gly Ala Ala Leu Gly Val
254                               580                               585                               590
256 (2) INFORMATION FOR SEQ ID NO: 3:
258   (i) SEQUENCE CHARACTERISTICS:
259       (A) LENGTH: 197 amino acids
260       (B) TYPE: amino acid
261       (C) STRANDEDNESS: Not Relevant
262       (D) TOPOLOGY: linear
264   (ii) MOLECULE TYPE: protein
269   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
271   Thr Leu His Arg Asn Glu Tyr Gly Ile Ala Ser Ile Leu Asp Ser Tyr
272   1                               5                               10                               15
274   Gln Cys Thr Ala Glu Ile Ser Leu Ala Asp Leu Ala Thr Ile Phe Phe
275   20                               25                               30
277   Ala Gln Phe Val Gln Glu Ala Thr Tyr Lys Glu Val Ser Lys Met Val
278   35                               40                               45
280   Lys Asp Ala Leu Thr Ala Ile Glu Lys Pro Thr Gly Asp Glu Gln Ser
281   50                               55                               60
283   Ser Gly Cys Leu Glu Asn Gln Leu Pro Ala Phe Leu Glu Glu Leu Cys
284   65                               70                               75                               80
286   His Glu Lys Glu Ile Leu Glu Lys Tyr Gly His Ser Asp Cys Cys Ser
287   85                               90                               95
289   Gln Ser Glu Glu Gly Arg His Asn Cys Phe Leu Ala His Lys Lys Pro
290   100                              105                              110
292   Thr Ala Ala Trp Ile Pro Leu Phe Gln Val Pro Glu Pro Val Thr Ser
293   115                              120                              125
295   Cys Glu Ala Tyr Glu Glu Asp Arg Glu Thr Phe Met Asn Lys Phe Ile
296   130                              135                              140
298   Tyr Glu Ile Ala Arg Arg His Pro Phe Leu Tyr Ala Pro Thr Ile Leu
299   145                              150                              155                              160
301   Leu Ser Ala Ala Gly Tyr Glu Lys Ile Ile Pro Ser Cys Cys Lys Ala
302   165                              170                              175
304   Glu Asn Ala Val Glu Cys Phe Gln Thr Lys Ala Ala Thr Val Thr Lys
305   180                              185                              190
307   Glu Leu Arg Glu Ser

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/08/879,469

DATE: 11/25/2003

TIME: 07:40:36

Input Set : N:\Crif3\RULE60\08879469.RAW.txt

Output Set: N:\CRF4\11252003\H879469.raw

```

308          195
310 (2) INFORMATION FOR SEQ ID NO: 4:
312   (i) SEQUENCE CHARACTERISTICS:
313       (A) LENGTH: 192 amino acids
314       (B) TYPE: amino acid
315       (C) STRANDEDNESS: Not Relevant
316       (D) TOPOLOGY: linear
318   (ii) MOLECULE TYPE: protein
323   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
325   Ser Leu Leu Asn Gln His Ala Cys Pro Val Met Lys Asn Phe Gly Thr
326   1          5          10          15
328   Arg Thr Phe Gln Ala Ile Thr Val Thr Lys Leu Ser Gln Lys Phe Thr
329           20          25          30
331   Lys Val Asn Phe Thr Glu Ile Gln Lys Leu Val Leu Asp Val Ala His
332           35          40          45
334   Val His Glu His Cys Cys Arg Ala Asp Val Leu Asp Cys Leu Gln Asp
335           50          55          60
337   Gly Glu Lys Ile Met Ser Tyr Ile Cys Ser Gln Gln Asp Thr Leu Ser
338           65          70          75          80
340   Asn Lys Ile Thr Glu Cys Cys Lys Leu Thr Thr Leu Glu Arg Gly Gln
341           85          90          95
343   Cys Ile Ile His Ala Glu Asn Asp Glu Lys Pro Glu Gly Leu Ser Pro
344           100         105         110
346   Asn Leu Asn Arg Phe Leu Gly Asp Arg Asp Phe Asn Gln Phe Ser Ser
347           115         120         125
349   Gly Glu Lys Asn Ile Phe Leu Ala Ser Phe Val His Glu Tyr Ser Arg
350           130         135         140
352   Arg His Pro Gln Leu Ala Val Ser Val Ile Leu Arg Val Ala Lys Gly
353           145         150         155         160
355   Tyr Gln Glu Leu Leu Glu Lys Cys Phe Gln Thr Glu Asn Pro Leu Glu
356           165         170         175
358   Cys Gln Asp Lys Gly Glu Glu Glu Leu Gln Lys Tyr Ile Gln Glu Ser
359           180         185         190
362 (2) INFORMATION FOR SEQ ID NO: 5:
364   (i) SEQUENCE CHARACTERISTICS:
365       (A) LENGTH: 201 amino acids
366       (B) TYPE: amino acid
367       (C) STRANDEDNESS: Not Relevant
368       (D) TOPOLOGY: linear
370   (ii) MOLECULE TYPE: protein
375   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
377   Gln Ala Leu Ala Lys Arg Ser Cys Gly Leu Phe Gln Lys Leu Gly Glu
378   1          5          10          15
380   Tyr Tyr Leu Gln Asn Glu Phe Leu Val Ala Tyr Thr Lys Lys Ala Pro
381           20          25          30
383   Gln Leu Thr Ser Ser Glu Leu Met Ala Ile Thr Arg Lys Met Ala Ala
384           35          40          45
386   Thr Ala Ala Thr Cys Cys Gln Leu Ser Glu Asp Lys Leu Leu Ala Cys
387           50          55          60

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/08/879,469

DATE: 11/25/2003

TIME: 07:40:37

Input Set : N:\Crf3\RULE60\08879469.RAW.txt

Output Set: N:\CRF4\11252003\H879469.raw

L:27 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:28 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:33 M:238 W: Alpha Fields not Ordered, Reordered [(C) CLASSIFICATION:] of (1)(vi)
L:55 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=1
L:682 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=9
L:699 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=10
L:716 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=11
L:733 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=12
L:750 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=13
L:767 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=14
L:784 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=15